Cell, Vol. 94, 147–150, July 24, 1998, Copyright ©1998 by Cell Press

Bacteriophage DNA Packaging: RNA Gears in a DNA Transport Machine

Minireview

Roger W. Hendrix
Pittsburgh Bacteriophage Institute and
Department of Biological Sciences
University of Pittsburgh
Pittsburgh, Pennsylvania 15260

The question of how the dsDNA bacteriophages get their genomic DNA on the inside of the virus particle, surrounded by their protein capsid on the outside, has been on the minds of virologists probably since Salvador Luria and Thomas Anderson made the first electron micrographs of phage T2 and saw tadpole-shaped objects that they likened to spermatozoa (Luria and Anderson, 1942). Initially it seemed obvious that the DNA must first condense into a compact form, after which the capsid proteins would come together to form a shell around it. The problem was framed in its current form about 25 years ago when it was shown—astonishingly at the time—that during the latter stages of a phage infection, an empty protein shell is assembled first, and then the DNA is somehow transported across the shell into the interior (Luftig et al., 1971; Kaiser et al., 1975). Studies aimed at understanding this process have had important technological benefits for the field—think of phage λ cosmid vectors and in vitro DNA packaging—but the DNA packaging process also has a number of features that recommend it as a model system for teasing out the secrets of some of the most central problems of modern molecular biology. These include how protein–protein interactions contribute to biological function, how protein–DNA interactions produce movement in a multicomponent structure, and more generally, how a complex biological machine is able to carry out an intricate task: what are all the molecular gears and levers, and how do they work together to carry out their role—in this case to pump the DNA up a steep thermodynamic gradient into the capsid? As described below, there are now clear indications that RNA has a specific and essential role in this process as well.

The phage biologists who have addressed these issues over the years have worked out a rather detailed picture of the overall process of DNA packaging (see, for example, Black, 1988). Briefly, the product of phage DNA replication, which may be a multigenome concatemer or a single-genome-sized molecule depending on the phage, interacts with a phage-encoded protein to initiate the packaging process. In the case of a typical concatemer-forming phage like the well-studied λ, the protein in question is the heterodimeric enzyme terminase, which binds site-specifically to the DNA and cuts it to form the first mature end of the soon-to-be-packaged DNA. The terminase stays with the DNA after cutting, and the terminase–DNA complex finds its way to a prohead—the empty protein shell that has formed in a separate assembly pathway—and binds. Insertion of the DNA into the prohead commences, cut end first, and continues until the head is full of DNA. During DNA insertion the terminase remains associated with the packaging complex and indeed probably forms an important component of the DNA transport machine. When it is time to make the second cut, the terminase is in place to do so, and after cutting finds itself bound to the DNA past the cut site, forming a new terminase–DNA complex that goes off to find another prohead and repeat the process.

On the prohead side of the packaging machine, the main actor in this play is the "portal" or "head-tail connector," which is an annular structure made of 12 copies (possibly 13 copies in some cases) of the phage-encoded portal protein (Carazo et al., 1986; Dube et al., 1993). The portal resembles a grommet, sitting on one of the 5-fold symmetric "corners" of the icosahedral capsid. There is a roughly DNA-sized hole in the middle of the portal, and in the mature phage this hole is lined up with the hole down the middle of the tail. When the DNA is injected into a new host, it must pass through the hole in the portal on its way down the tail and into the cell. For DNA packaging (which happens before the tail is on the scene) most models envision the DNA entering the capsid through the hole in the portal as well, though there is neither universal agreement nor direct evidence on this point. The portal is thought to be the place where the terminase binds during packaging, and the terminase–portal complex, with the DNA passing through it, is generally considered to be the heart of the machine that moves the DNA. The terminase is an ATPase and therefore most likely the first molecule in the chain of energy transduction that leads to DNA transport. A remarkable ability of the DNA packaging machine is to sense how much DNA has been packaged into the head and "decide" when the head is full and it is time to trigger the terminal cleavage of the DNA. Genetic experiments implicate the portal in this process (Casjens et al., 1992; Tavares et al., 1992).

One of the technical advances that has made it possible to learn much of what we know about DNA packaging is the development of in vitro packaging systems. Such systems have been worked out for a number of phages, and as they have become better defined and more efficient it has been possible to show, for example, that DNA packaging is driven by ATP hydrolysis to the tune of about one ATP per two base pairs translocated into the capsid (Guo et al., 1987b). One of the best such in vitro packaging systems is that of the *Bacillus subtilis* phage ϕ29, and it was with this system that the startling discovery was reported several years ago that there is an absolute requirement for presence of an RNA molecule in order to see packaging (Guo et al., 1987a). The RNA in question, termed "packaging RNA" or "pRNA," is a 174-nucleotide molecule encoded in the phage genome. A highly folded secondary structure initially postulated from sequence gazing has been largely supported by mutational analysis and by phylogenetic analysis of the apparently homologous pRNA sequences from a number of phages closely related to ϕ29. The pRNA is found associated with the prohead in about six copies, but it is completely absent from the mature virion, and in that sense can be considered an essential chaperone or facilitator of the packaging event. These

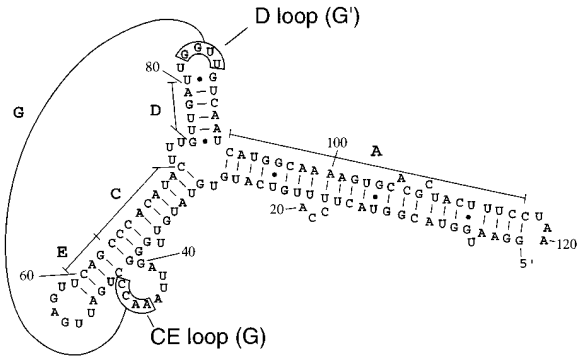

Figure 1. Sequence and Predicted Secondary Structure of the ϕ29 pRNA

A ring of six of these pRNAs is proposed to bind to the portal. The shading shows the portion of the sequence implicated in portal binding; the indicated loops are the binding sites between adjacent pRNAs. The sequence shown is the 120-base truncated form of the pRNA, which has the full activity of the 174-base form and is used in the experiments described here (from Zhang et al. 1998, with permission).

RNAs have generated substantial interest, first because they promise to contribute to the expanding catalog of previously unexpected biochemical feats that RNA is capable of, and second, because they may be the missing piece of the puzzle for understanding the mechanism by which DNA is transported into the phage capsid.

The latest attempts to tame the pRNAs are described in a pair of papers from the laboratories of Dwight Anderson and Peixuan Guo that appear in this month's issue of *Molecular Cell* (Guo et al., 1998; Zhang et al., 1998). These authors report elegant sets of experiments in which they use a combination of the ϕ29 in vitro DNA packaging assay and constructed mutants of the pRNA to count the number of pRNA molecules in the packaging-competent proheads and to infer their topology.

The take-off point for the experiments presented by both groups is an indication from earlier studies that two of the single-stranded loops on the pRNA interact with each other, apparently through complementary Watson-Crick base pairing. Since the in vitro packaging reaction is absolutely dependent on the pRNA, it can be used to test the functional importance of this interaction. Changing the sequence in either of the loops ("G" and "G′" in Figure 1) kills packaging activity, but if both mutant sequences are put in the same molecule—and if the mutant sequences are complementary—packaging is fully restored. This confirms the importance of the interaction. Things get interesting when a mixture of the two individually mutant pRNAs is supplied to the reaction. Even though each mutant by itself is inactive, the mixture is as active as wild type. The interpretation of this is that the interaction between loops is not between the two loops on one pRNA molecule but between complementary loops on different molecules. Thus, the mutant G loop on one pRNA molecule is complementary to the mutant G′ loop on the other, and the wild-type G loop on the second pRNA is complementary to the wild-type G′ on the first. Thus, the complementary loops serve as linkers in a chain of pRNAs. If we make the reasonable assumption (as the authors do) that the chain of pRNAs forms a closed structure (i.e., a ring), then this result with alternately complementary pRNAs argues that the number of pRNAs in the active structure is a multiple of two. A similar experiment with three different serially complementary pRNAs argues that the active structure has a multiple of three pRNAs, and this together with the previous experiments argues that the number must be six or a multiple. Guo et al. (1998) even make a set of six different mutants that should work only as a set of 6N molecules, and they measure full activity.

Both groups provide independent sorts of evidence that argue for pRNA hexamers as the active form. Guo's group had argued earlier (Trottier and Guo, 1997) that when a nonfunctional mutant pRNA is mixed with wild type, a single mutant pRNA in the packaging complex is enough to completely kill activity. This allows them to calculate how fast bulk packaging activity should decrease as the reaction is doped with defective pRNA, and the data they show in the present paper fit very nicely with the prediction for a hexamer structure and are inconsistent with the prediction for a dodecamer. Both groups show evidence for pRNA oligomers in nondenaturing polyacrylamide gels (but only when the molecules have the appropriate complementarities). The oligomeric forms are identified as dimers and hexamers, and Zhang et al. (1998) confirm this by measuring the complex molecular weights using sedimentation equilibrium.

Aside from mechanistic questions about how the pRNA carries out its role in DNA packaging, the circular hexamer of pRNAs implied by these experiments is something new and interesting in the world of RNA structure. These experiments show clearly that the specificity of interaction between pRNAs comes from the apparent base-pairing interactions between the loops and is in fact due primarily to just two crucial base pairs. Presumably other interactions between the pRNAs are required to give the hexamers the observed stability, and what these are and how their formation interacts with forming the base pairs between loops remain for future studies. However, Zhang et al. (1998) fail (appropriately, I would say) to resist the temptation to calculate a three-dimensional model of the pRNA hexamer, and this model is shown in their paper. Although it seems very unlikely that the model, at the current state of knowledge, is correct in all its details, it does have one interesting property that would likely be true of most models of this general sort—namely, that the hole in the middle of the donut is just the right size to slide over the outer end of the portal where it protrudes from the capsid shell. Together with the fact from previous work that pRNA binds to the portal, this suggests the position for the pRNA hexamer on the prohead indicated in Figure 2. Recent progress in reconstructing images of ϕ29 and its precursor structures from cryo-electron microscopic data may allow a direct test of this structural hypothesis soon.

The biggest hole in our understanding of how DNA is packaged by ϕ29 or any of the other dsDNA phages remains the actual mechanism by which DNA is transported across the capsid boundary—and this is not for

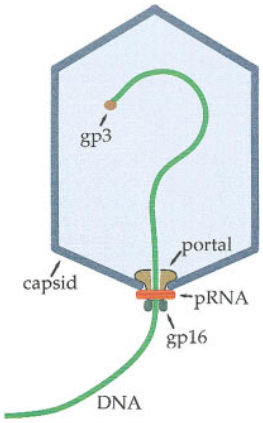

Figure 2. Schematic Representation of the DNA Packaging Machine of Bacteriophage ϕ29

want of trying by workers in the field. I committed a model to the literature myself in 1978 (Hendrix, 1978) proposing that the portal rotates with respect to the rest of the capsid as the helical DNA passes through it to fill the capsid. The fact that portal rotation is still a prominent component of discussions of packaging mechanisms 20 years later may have less to do with the virtues of the idea than with the difficulty of ruling it out experimentally. The more radical version of the portal rotation model holds that the energy from ATP hydrolysis is coupled directly to rotation of the portal, and that the DNA is screwed into the head as a consequence of its "threaded" helical structure, much as a bolt moves through its nut when the nut is turned with a wrench. Guo et al. (1998) favor such a model, in which they imagine that branches of the pRNAs reach out and act as arms or oars that interact sequentially with parts of the capsid and impart increments of rotational motion to the portal–pRNA complex. Such a mechanism, if correct, would clearly expand our understanding of what RNA can do, but working out the mechanism of energy transduction will be challenging. A more conventional mechanism would propose that ATP hydrolysis is used to apply force between the portal–terminase–pRNA complex and the DNA molecule and thereby cause that complex to walk down the DNA. The idea that the donut-shaped DNA packaging complex might be driven along the DNA by ATP hydrolysis is reminiscent of current views on how donut-shaped helicases such as phage T7 gene 4 protein and *E. coli* DnaB work (Baker and Bell, 1998), and it may be worth considering whether this parallel illuminates the DNA packaging mechanism. Since the packaging complex is part of the capsid shell, walking it down the DNA is equivalent to transporting DNA into the capsid. In such a model we might still expect the portal and the associated components to rotate passively to relieve the twist that would otherwise accumulate in the DNA, assuming the portal complex follows the DNA helix as it moves along it. Two possible roles for the pRNA seem evident in this case: it might be a working part of the machine that moves the portal along the DNA or it might have a more structural role—for example, providing a bearing surface to facilitate portal rotation. Other models of DNA packaging have been proposed, some more complex than these and not all incorporating the idea of portal rotation, and I will not enumerate them here. Suffice it to say that when we understand the biochemical role of pRNA in DNA packaging, it is likely that we will be well on the way to understanding the whole mechanism.

As more is learned about virion assembly in dsDNA phages, both from direct studies of virion structure and assembly and from comparisons of genomic sequences and organization, it is becoming quite clear that these phages share common ancestry, even when all vestiges of evidence for similarity have been lost from the nucleotide sequences of their genomes and the amino acid sequences of their structural proteins. (There is even getting to be good reason to believe that some of the animal viruses—notably Herpesvirus and Adenovirus—are cousins of the dsDNA phages, hailing from different branches of the same phylogenetic tree.) What we observe when we investigate virion assembly of the dsDNA phages is that the overall strategy of assembly is strongly conserved from one phage to another but that some of the less fundamental details can be quite variable. Comparing the ways that different phages accomplish similar goals has frequently illuminated not only which are the fundamental parts of the assembly process and which variations on a theme, but sometimes even what the real functions are. Comparing ϕ29 to other phages, it shares the property of packaging its DNA in linear order into a preformed shell through a process that uses ATP hydrolysis for energy and a structurally conserved portal as the entry point for the DNA. ϕ29 differs from many other phages in that the product of its DNA replication is not a multigenome concatemer but a unit length genome with proteins covalently bound to the 5′ ends of the DNA. Given this, it does not need a terminase enzyme to create the ends of the genomic DNA, but ϕ29 does have an essential DNA packaging protein, gp16, a pRNA-stimulated ATPase (Grimes and Anderson, 1990) that plausibly carries out the other roles that terminase is responsible for in other phages. In this context, we can ask whether the use of pRNA in DNA packaging is universal among the dsDNA phages. To date, pRNA has been identified only in ϕ29 and its close relatives, despite efforts to find equivalent RNAs in other phage systems. This may simply reflect a technical difficulty in detecting pRNAs in other systems. On the other hand, it may well turn out that some phages do not have pRNAs. Experience with other aspects of virion assembly argues that the essential function that pRNA accomplishes for ϕ29 will still be accomplished in other phages, even if not by a pRNA. Thus, I expect to see one of two outcomes: either all dsDNA phages will be shown to encode pRNAs that carry out the same function as do the ϕ29 pRNAs (whatever that may turn out to be), or that all these phages will accomplish that function, but some will do so by a biochemically distinct means. I can't decide which outcome will be the more interesting.

Selected Reading

Baker, T.A., and Bell, S.P. (1998). Cell *92*, 295–305.

Black, L.W. (1988). In The Bacteriophages, vol. 2, R. Calendar, ed. (New York: Plenum Press), pp. 321–373.

Carazo, J.M., Donate, L.E., Herranz, L., Secilla, J.P., and Carrascosa, J.L. (1986). J. Mol. Biol. *192*, 853–867.

Casjens, S., Wyckoff, E., Hayden, M., Sampson, L., Eppler, K., Randall, S., Moreno, E.T., and Serwer, P. (1992). J. Mol. Biol. *224*, 1055–1074.

Dube, P., Tavares, P., Lurz, R., and van Heel, M. (1993). EMBO J. *12*, 1303–1309.

Grimes, S., and Anderson, D. (1990). J. Mol. Biol. *215*, 559–566.

Guo, P., Erickson, S., and Anderson, D.L. (1987a). Science *236*, 690–694.

Guo, P., Peterson, C., and Anderson, D. (1987b). J. Mol. Biol. *197*, 229–236.

Guo, P., Zhang, C., Chen, C., Garver, K., and Trottier, M. (1998). Mol. Cell *2*, 149–155.

Hendrix, R.W. (1978). Proc. Natl. Acad. Sci. USA *75*, 4779–4783.

Kaiser, A.D., Syvannen, M., and Masuda, T. (1975). J. Mol. Biol. *91*, 175–186.

Luftig, R., Wood, W., and Okinada, R. (1971). J. Mol. Biol. *57*, 555–573.

Luria, S.E., and Anderson, T.F. (1942). Proc. Natl. Acad. Sci. USA *28*, 127–130.

Tavares, P., Santos, M.A., Lurz, R., Morelli, G., de Lencastre, H., and Trautner, T.A. (1992). J. Mol. Biol. *225*, 81–92.

Trottier, M., and Guo, P. (1997). J. Virol. *71*, 487–494.

Zhang, F., Lemieux, S., Wu, X., St.-Arnaud, D., McMurray, C.T., Major, F., and Anderson, D. (1998). Mol. Cell *2*, 141–147.